Assessing a New Coupled Data Assimilation System Based on the Met Office Coupled Atmosphere–Land–Ocean–Sea Ice Model

Met Office, Exeter, United Kingdom

(Manuscript received 10 April 2015, in final form 5 August 2015)

ABSTRACT
A new coupled data assimilation (DA) system developed with the aim of improving the initialization of coupled forecasts for various time ranges from short range out to seasonal is introduced. The implementation here is based on a “weakly” coupled data assimilation approach whereby the coupled model is used to provide background information for separate ocean–sea ice and atmosphere–land analyses. The increments generated from these separate analyses are then added back into the coupled model. This is different from the existing Met Office system for initializing coupled forecasts, which uses ocean and atmosphere analyses that have been generated independently using the FOAM ocean data assimilation system and NWP atmosphere assimilation systems, respectively. A set of trials has been run to investigate the impact of the weakly coupled data assimilation on the analysis, and on the coupled forecast skill out to 5–10 days. The analyses and forecasts have been assessed by comparing them to observations and by examining differences in the model fields. Encouragingly for this new system, both ocean and atmospheric assessments show the analyses and coupled forecasts produced using coupled DA to be very similar to those produced using separate ocean–atmosphere data assimilation. This work has the benefit of highlighting some aspects on which to focus to improve the coupled DA results. In particular, improving the modeling and data assimilation of the diurnal SST variation and the river runoff should be examined.

1. Introduction
Forecasting systems for short-range weather and ocean prediction have been run separately at the Met Office for many years with the weather forecasts using prescribed ocean surface temperatures and sea ice fields, and with the ocean forecasts using atmospheric forcing fields from the Met Office’s numerical weather prediction (NWP) system. It has long been known that coupling between the various earth system components (the ocean, atmosphere, sea ice, and land) produces improved forecasts on seasonal and longer time scales (e.g., Neelin et al. 1994). Recent research by Johns et al. (2012) and Johns et al. (2015, manuscript submitted to Mon. Wea. Rev.) also shows significant improvements on shorter time scales (out to 15 days) to both atmospheric and ocean forecast skill in the tropics when using a coupled model as opposed to either atmosphere or ocean models run separately. There is an increasing interest within the GODAE Oceanview community in coupled predictions on shorter time scales (Brassington et al. 2015). Pellerin et al. (2004) demonstrate significant forecast improvements associated with the changes in sea ice cover when using a coupled atmosphere–ocean–sea ice model in the Gulf of St. Lawrence. Improvements to the modeling of tropical cyclones were shown by Chen et al. (2010) when using coupled models.

For the development of coupled forecasting at medium range, and for operational seasonal forecasts [e.g., the Global Seasonal forecast system version 5 (Glosea5) by MacLachlan et al. (2015)], the need for accurate initial conditions has been met by the use of separately generated atmosphere–land and ocean–sea ice state estimates. The system used to generate the atmospheric initial conditions at the Met Office has been the atmospheric 4DVAR data assimilation (DA) system of Rawlins et al. (2007), run at the operational NWP resolution (which is higher resolution than is used for existing coupled applications). The ocean initial conditions are currently generated using a 3DVAR scheme as part of the FOAM system (Waters et al. 2014; Blockley et al. 2014). Coupled seasonal forecasts are then initialized by

Corresponding author address: Daniel Lea, Met Office, FitzRoy Road, Exeter EX1 3PB, United Kingdom.
E-mail: daniel.lea@metoffice.gov.uk

DOI: 10.1175/MWR-D-15-0174.1
reconfiguring the atmospheric state to the resolution of the atmospheric part of the coupled model, and using these with the ocean initial conditions. As the initial conditions come from separate systems these are quite likely to contain inconsistencies, potentially leading to imbalances in the coupled model.

The work presented here is a first step toward being able to provide more consistent initial conditions for coupled ocean–sea ice–atmosphere–land forecasts at all time ranges. We would like to understand whether such initial conditions will provide improved forecasts for the various components of the system. An important by-product of such a system is the ability to easily diagnose the model errors that grow in the first few hours of the coupled forecast through the changes determined by the data assimilation, which should enable coupled models to be improved for all applications.

Coupled data assimilation is a relatively new area of research and involves a number of challenges. Laloyaux et al. (2014, 2015) have recently developed a weakly coupled DA system, designed for coupled climate re-analyses, with a relatively low-resolution (1°) ocean component. The approach taken here is to develop a system based on the existing coupled model and the separate data assimilation systems for ocean, sea ice, land, and atmosphere introduced above. The coupled model provides background information for separate analyses in each subcomponent, with the increments being added back into the coupled model. We use existing estimates of forecast error covariances with no attempt to model coupled error covariances. This approach is called weakly coupled data assimilation in, for example, Brassington et al. (2015) and later in this paper. A scheme to account for systematic model errors in the ocean is included based on the scheme of Bell et al. (2004) and Balmaseda et al. (2007). The mismatch of time scales between the ocean and atmosphere is a major difficulty, as the two data assimilation systems are constrained to run using the same time window in the weakly coupled DA approach employed here. We have taken the approach of using the shorter (6h) atmospheric time window for both components, as this is more appropriate for operational coupled NWP. It may also have required significant developments to the Met Office atmospheric data assimilation system to allow a longer time window.

The approach of developing a weakly coupled data assimilation system is theoretically rather simple, even though it is technically challenging. It is also a necessary first step to developing more fully coupled data assimilation methods in which observations from one component (e.g., the ocean) may directly affect another component (e.g., the atmosphere). For example, we can use runs with this system to estimate biases and covariances, which may then be applied to more sophisticated coupled data assimilation techniques. The subject of this paper, however, is focused on assessing the performance of the weakly coupled DA system and identifying aspects of the system that may be improved in future. Assessing the coupled error covariances will be left to future work.

In section 2 we describe the coupled model components and the data assimilation systems used in more detail. In section 3 the results of a 13-month run of the coupled DA system are assessed. In section 4 the impact of coupled DA on coupled forecasts is considered. Finally, section 5 contains the discussion and our conclusions.

2. System description

In this section we briefly describe the implementation of the data assimilation systems within the weakly coupled data assimilation framework, focusing mainly on the differences among the existing implementations of the various components. The experimental framework adopted to test the system is then outlined.

a. Coupled model

The coupled model in our weakly coupled DA system combines the Met Office Unified Model atmosphere (MetUM; Davies et al. 2005) and the Joint U.K. Land Environment Simulator land surface model (JULES; Best et al. 2011) coupled to the Nucleus for European Modeling of the Ocean model (NEMO; Madec 2008) and the Los Alamos sea ice model (CICE; Hunke and Lipscomb 2010). The MetUM and JULES models use the Global Atmosphere 4.0 and Global Land 4.0 science configurations (Walters et al. 2014) run at N216 horizontal resolution (432 × 325 grid points, ~60 km in the midlatitudes). The MetUM uses 85 vertical levels, with 50 levels below 18 km (and hence at least sometimes in the troposphere), 35 levels above this (and hence solely in or above the stratosphere), and a fixed model lid 85 km from the surface. JULES comprises four soil layers and uses a tiled approach to its surface exchange with each land point subdivided into five types of vegetation (broadleaf trees, needleleaved trees, temperate C3 grass, tropical C4 grass, and shrubs) and four non-vegetated surface types (urban areas, inland water, bare soil, and land ice). NEMO and CICE are run in an ORCA025 configuration described in Blockley et al. (2014). This uses a tripolar grid with roughly 1/4° horizontal resolution. NEMO uses 75 vertical levels with approximately 1-m resolution in the top 10 m of the ocean, with the resolution decreasing in the deeper ocean. CICE is run with five ice thickness categories.
The atmosphere–land and ocean–sea ice components are coupled at hourly intervals using the OASIS coupler (Valcke 2006) as described in Hewitt et al. (2011). The initial conditions of the coupled model are corrected using two separate 6-h window data assimilation systems: a 4DVAR system for the atmosphere with associated soil moisture content nudging and snow analysis schemes, and a 3DVAR–first-guess-at-appropriate-time (FGAT) system for the ocean and sea ice. The background information for all of the DA systems comes from a previous 6-h forecast of the coupled model.

b. Atmosphere data assimilation system

The atmosphere data assimilation scheme is an incremental strong constraint 4DVAR system similar to that described in Rawlins et al. (2007). The scheme implemented here has climatological background error covariances, unlike the now operational hybrid-VAR system that combines climatological covariances with those generated from an ensemble to represent errors specific to the synoptic situation. An incremental system involves running one or more outer-loop iterations where the full nonlinear state is updated (Courtier 1994). Each time this occurs, there is an inner loop where a cost function linearized about the full nonlinear state is minimized. Running multiple outer loops is expensive so most systems run only a few. The Met Office 4DVAR system, while capable of running multiple outer loops, currently uses a single outer loop. ECMWF IFS, on the other hand, has two outer-loop cycles for lower-resolution atmospheric analyses and four outer loops for higher-resolution systems.

In the Met Office system, at each cycle, the model is run for 12 h, where the first 6 h cover the window of the current assimilation cycle, including the addition of the assimilation increments, and the last 6 h produce background trajectories for the next assimilation cycle. The 6-h assimilation windows are run from 0300 to 0900, 0900 to 1500, 1500 to 2100, and 2100 to 0300 UTC and an analysis is produced in the middle of each window. The 0000 and 1200 UTC analyses each day are used to initialize the forecasts.

The atmosphere observations are extracted from the Met Office’s observational database. Different types of satellite and in situ observations are assimilated, including temperature, wind, humidity, pressure, and direct radiances. A quality control check is first performed to remove any unrealistic observations [see Rawlins et al. (2007) for more details]. The observations are then compared to the trajectory produced by the nonlinear model and innovations (observation and equivalent model values) are generated at the observation positions.

The inner loop uses a perturbation forecast (PF) model run at a coarser horizontal resolution of N108 (216 × 163 grid points, ~120 km in the midlatitudes) but the same 85 levels in the vertical. This model is a linearization of the MetUM model about a full-resolution background trajectory of the MetUM model. A simplification operator and its generalized inverse allow us to swap from one model to the other. Besides decreasing resolution, the simplification operator also consolidates multiple moisture and cloud variables into a single variable (Lorenc 2003). The background error covariance has been calculated using the ensemble method described in Clayton et al. (2012). The result is increments of the potential temperature, zonal and meridional components of wind, total mass of water, density, and Exner pressure. These increments are provided to the MetUM model, which applies them before starting its first time step.

For the land surface we produce a separate 3DVAR analysis of screen temperature and humidity, which is then solely used to correct the soil moisture via the scheme of Best and Maisey (2002). The assumption of the scheme is that in certain conditions (e.g., during the daytime and where wind speeds are less than 10 m s\(^{-1}\)) the dominant cause of errors in surface temperature and humidity is errors in the model soil moisture. The soil moisture is then further improved by the method of Dharssi et al. (2011) using Advanced Scatterometer (ASCAT) observations of soil wetness. The resulting soil moisture field is then applied to the model in the middle of the assimilation window. Note there is no direct correction of the land surface temperature by assimilation; changes in the modeled land surface temperature occur instead indirectly via changes in the soil moisture and atmospheric state. The atmospheric state itself is corrected in the 4DVAR atmosphere assimilation described above.

Once a day, a snow analysis is performed using the daily map of Northern Hemisphere snow cover data from the Interactive Multisensor Snow and Ice Mapping System (IMS) from the National Environmental Satellite, Data, and Information Service (NESDIS). The snow cover observations are converted into a fractional cover product on the model grid. An analysis is then produced by combining these data with the snow amount field of the MetUM. The addition of snow in particular is done by relating the snow-water equivalent to the fractional cover (e.g., Drusch et al. 2004) and using result in the specification of the albedo. The snow amount analysis is then directly inserted into the model initial conditions.

In the atmosphere control run, the sea surface temperature (SST) and the sea ice fraction are updated
once a day with the Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA; Donlon et al. 2012). The data are interpolated onto the MetUM grid before being read into the model. The SST is persisted during the four cycles of the day. In the coupled run, by contrast, the SST is taken, via OASIS, from the top level of the ocean and the sea ice fraction from the sea ice model. Details on the initialization of SST, sea ice, and other aspects of the ocean system are described in the next section.

c. Ocean system

The data assimilation scheme used for the ocean and sea ice is the variational data assimilation scheme of NEMO (NEMOVAR; Mogensen et al. 2009, 2012), which is run in 3DVAR-FGAT mode. The implementation here is based on that of the FOAM system (Waters et al. 2014). The scheme assimilates in situ and satellite SST data, satellite altimeter sea level anomaly (SLA) data, satellite sea ice concentration data, and in situ temperature and salinity profiles from various sources including Argo, moored buoys, and temperature profiles from XBTs and marine mammals. Increments are applied to the model using the Incremental Analysis Update (IAU) scheme of Bloom et al. (1996). Combining the FGAT scheme and IAU requires an observation operator run (of 24 h in FOAM), where the observations are compared to the model at the correct time followed by an IAU run over the same time period where increments calculated from those observations are applied to the model.

A number of changes to the FOAM data assimilation setup were required in order to implement it within the weakly coupled data assimilation framework. The first one was purely a technical change in which the order of the assimilation cycling was made to match the atmospheric cycling. FOAM currently performs its observation operator and IAU in separate NEMO runs; this was altered so that IAU and the observation operator could be run in one combined NEMO model run. The second major change was to make the time window of the ocean match the atmosphere time window of 6 h. This required several changes to the NEMO code (which has not previously been run on subdaily time scales to our knowledge) to make sure surface fluxes were applied at the correct time and that the restart time of the model was correct. We should also retune the error covariances because of the change in the time window. However, because errors do not grow linearly, there is no simple scaling that can be applied so this work is left for the future, when we will do a full reestimation of the error covariances in the coupled model (probably using the output of the runs presented here). The third change was to the period over which increments are added to the model in the IAU. Forecasts were to be run from the center of each assimilation time window, and so the increments need to be fully added into the model by that time to ensure that the forecasts are properly initialized. The IAU period was therefore reduced from 24 to 3 h for the experiments described here.

The impact of the above assimilation changes to FOAM DA was tested by running two ocean-only experiments for 1 month: one with the standard FOAM 24-h time window and 24-h IAU, and one with the 6-h time window with 3-h IAU. Innovation (observation minus model background) statistics (Table 1) from the two runs indicate the impact from these changes on the temperature and salinity profile statistics, and on the SLA statistics, to be rather small. For SST, the new cycling results in a reduction of the errors; this is likely because there is still a large amount of SST data available in the shorter time window, but there is less time for errors to grow. Overall, we do not see any serious problems arising from the switch to 6-h cycling.

d. Experimental setup

Previous experiments (e.g., Johns et al. 2012; Johns et al. 2015, manuscript submitted to Mon. Wea. Rev.) have shown the impact of coupled models compared with uncoupled ocean or atmosphere models on medium-range forecasts. Here, the aim is not to repeat those experiments but to focus on the impact of the coupled initialization strategy on the performance of the data assimilation and the short-range coupled forecasts.

Three data assimilation runs were performed. The first one (cpld_da) produces ocean–atmosphere–land–sea ice analyses using the weakly coupled data assimilation system described above. An atmosphere–land–only data assimilation experiment (actl_da) was run, with the SSTs and sea ice coming from OSTIA (an independent SST analysis used operationally for NWP). Fluxes from this run were then used to force an ocean–sea-ice-only data assimilation run (octl_da). To isolate the impact of coupled DA, the idea is for the atmosphere and ocean control runs to replicate the existing operational separate ocean and atmosphere assimilation

<table>
<thead>
<tr>
<th>Observation Type</th>
<th>6 h</th>
<th>24 h</th>
</tr>
</thead>
<tbody>
<tr>
<td>SST in situ (°C)</td>
<td>0.409</td>
<td>0.524</td>
</tr>
<tr>
<td>SLA (m)</td>
<td>0.078</td>
<td>0.083</td>
</tr>
<tr>
<td>Sea ice concentration</td>
<td>0.035</td>
<td>0.035</td>
</tr>
<tr>
<td>T profile (°C)</td>
<td>0.635</td>
<td>0.648</td>
</tr>
<tr>
<td>S profile (psu)</td>
<td>0.137</td>
<td>0.148</td>
</tr>
</tbody>
</table>
systems but as far as possible by using the same model configurations as the coupled model.

Table 2 gives an overview of the experiments that have been run. The data assimilation experiments were run over a 13-month period (1 December 2011–31 December 2012). Additionally, coupled forecasts were initiated from the coupled DA analyses (fc_cpld) and from the separate ocean and atmosphere analyses (fc_ctl). These were started every 12 h (at 0000 and 1200 UTC each day) for three periods: 1–31 December 2011, 26 August–15 September 2012, and 20–31 October 2012. The December set of 5-day forecasts was run to investigate the initial impact of coupled DA, while the August–September and October sets were longer 10-day forecasts originally chosen to correspond to the peak of the Indian monsoon and to the peak of the North Atlantic hurricane season, respectively. We primarily assess the forecasts by comparing them to observations. More detailed case studies are left for future work.

The data assimilation experiments are assessed in the following sections by examining the observation-minus-background (innovation) statistics for the oceans and atmosphere, differences between analysis fields, and by looking at the mean data assimilation increments. Results of the analysis are discussed in the next section. Then, in section 4, we show some results of the forecast experiments.

3. Analysis results

a. Atmosphere assessment

A straightforward method of assessing assimilation system performance is to compare the model output to observations. Table 3 shows some summary RMS observation and model background differences. The results for both systems are rather similar. In the Southern Hemisphere and in the tropics the coupled DA performs slightly better in surface pressure, but these differences between the runs do not appear to be statistically significant. Also, the mean innovations (not shown) show no significant differences between the coupled and control DA simulations.

It is interesting also to compare the results to standard Met Office NWP (also in Table 3). This shows that the system is generally working. The results are expected to be degraded compared to the NWP atmospheric DA system because this approach used an atmosphere at a higher resolution (N512).

Another way of looking at the impact of coupled DA is to examine the model fields. This will allow us to see differences that may not be sampled by observations. The impact of the coupled ocean on the atmosphere can be seen in Fig. 1, for example, which shows that the ocean currents particularly in the tropics affect the wind. The ocean currents are quite similar in the control and coupled cases, but in the control the atmosphere model, uncoupled from the ocean, is not affected by the ocean currents. The impact of the coupling is seen even though the atmosphere in both cases is assimilating the same wind data (primarily from scatterometer over the ocean). Whether the winds are improved (compared to

<table>
<thead>
<tr>
<th>Expt</th>
<th>Description</th>
<th>Periods</th>
</tr>
</thead>
<tbody>
<tr>
<td>actl_da</td>
<td>Atmosphere-only DA, OSTIA SST, and sea ice</td>
<td>1 Dec 2011–31 Dec 2012</td>
</tr>
<tr>
<td>octl_da</td>
<td>Ocean-only DA with fluxes from actl_da</td>
<td>1 Dec 2011–31 Dec 2012</td>
</tr>
</tbody>
</table>

* Forecasts (fc_cpld and fc_ctl) of 5 or 10 days run every 12 hours over the periods indicated.
the observations) is not clear; nevertheless, it does indicate the potential of coupled DA.

### b. Ocean assessment

As in the atmosphere assessment, we begin by assessing the ocean model compared to observations. Table 4 shows innovation (observation minus background) statistics using the observation feedback files produced by NEMO. We have included FOAM statistics from the reanalysis in Blockley et al. (2014). As the time window for the assimilation differs in the coupled and control DA experiments and the time period is different (December 2011–November 2012 for FOAM), this is only a rough comparison, but it indicates that the ocean assimilation is generally functioning well in the experiments. Additionally, the coupled and control DA experiments have generally similar performance, with the profile temperature and salinity results being almost identical. The Kara et al. (2000) mixed layer depth is slightly improved by the coupled DA. However, there is some degradation in the coupled SST and SLA statistics compared to the ocean control results of 7% and 2%, respectively. The SST case in particular merits further investigation because of its importance as the interface between the atmosphere and ocean in the coupled model.

To investigate the small degradation in the SST statistics of the coupled ocean compared to ocean control, we focus on one region, the South Pacific, which shows the biggest differences in RMS SST innovations in December. Figure 2 shows the observation minus coupled and control ocean model background results following a drifter (ID 32546) in the South Pacific in the region of 24°–21°S and 140°–138°W. Both the ocean control and the coupled analysis show a diurnal cycle of errors. In particular, when the observed diurnal cycle is strongest because of light winds and clear skies (e.g., 16 December 2011), both models’ temperatures are in excess of the observed value in the peak of the day, but at night the situation is reversed and both models are somewhat colder than the observations. The coupled model errors compared to observations is somewhat amplified compared to the ocean control, which is largely what leads to increased RMS SST innovations. It is important to note that the ocean model has a 1-m-thick top box and no explicit diurnal skin model and therefore is not well equipped to correctly model the diurnal cycle either at the surface skin or at the 30–50-cm drifter depth. Kawai and Wada (2007) give some examples of the detailed vertical structure in the top 1 m, which will not be resolved in our ocean model. Work is in progress (Shelly et al. 2015) to include an explicit diurnal model that should improve the performance of the coupled ocean

<table>
<thead>
<tr>
<th></th>
<th>Coupled</th>
<th>Ocean control</th>
<th>FOAM*</th>
</tr>
</thead>
<tbody>
<tr>
<td>SST in situ GO (°C)</td>
<td>0.4147</td>
<td>0.3984</td>
<td>0.4747</td>
</tr>
<tr>
<td>SST in situ TP (°C)</td>
<td>0.3212</td>
<td>0.2990</td>
<td>0.3108</td>
</tr>
<tr>
<td>SLA GO (m)</td>
<td>0.0746</td>
<td>0.0730</td>
<td>0.0728</td>
</tr>
<tr>
<td>Sea ice concentration GO</td>
<td>0.0296</td>
<td>0.0295</td>
<td>0.0288</td>
</tr>
<tr>
<td>Mixed layer depth GO (m)</td>
<td>35.9741</td>
<td>36.3326</td>
<td>34.5567</td>
</tr>
<tr>
<td>Temp profile GO (°C)</td>
<td>0.6250</td>
<td>0.6199</td>
<td>0.6198</td>
</tr>
<tr>
<td>Salinity profile GO (psu)</td>
<td>0.1243</td>
<td>0.1243</td>
<td>0.1308</td>
</tr>
</tbody>
</table>

*Note the FOAM reanalysis statistics are shown as a rough comparison as the time window is 24 h compared to 6 h for the coupled and ocean control experiments. Additionally, the FOAM statistics are for the period December 2011–November 2012 whereas other statistics are for December 2011–December 2012.

![Fig. 1.](image-url)
model. Meanwhile, we do not consider this to be a specific serious issue for coupled DA; rather, it is an issue with the coupled model. Indeed, later in section 4 we show that the difference between the control and coupled DA SSTs quickly disappears in the coupled forecasts.

In Table 4 we also note that the SLA statistics are somewhat degraded in the coupled DA run compared to the control. When the innovations are spatially binned, it is clear that the degradation in the coupled model is very localized. Figure 3 shows the increase in RMS occurs only in certain shelf regions and major river outflows. The altimeter data we receive are tidally and barotropically filtered, and while the ocean model does not include tidal forcing, it may not be appropriate to compare the model and data in shelf regions. The increase in SLA error in the shelf regions is therefore not considered to be concerning. This will, however, need to be investigated for any future coupled DA work using higher-resolution shelf models. In these cases we may also assimilate more suitable unfiltered SLA data, for example, the Tailored Altimetry Products for Assimilation Systems (TAPAS) data from AVISO.

Figure 2. Diurnal cycle representation of the coupled DA analysis and the control ocean DA analysis of a drifting float (ID 32546) in the South Pacific during December 2011. (a) The observed temperature and the temperature for the two experiments at the observation location. (b) The surface wind and (c) cloud fraction interpolated to the observation location.
The salinity increase near the eastern seaboard of South America may be more of an issue. Looking in particular at the Rio de la Plata region at around 30°S, 50°W associated with the increased SLA RMS, there is a local degradation in the fit to salinity data in the same region (Fig. 4). Large differences in the model salinity are also evident at the ends of the runs in Fig. 5. Significant differences in salinity are evident around many coastal regions: a large plume (of salinity difference) is evident emanating from the Rio de la Plata. There is no coupled feedback in surface salinity (unlike SST) and so without data assimilation errors can build with time. It should be noted that there is, however, a weak Haney surface salinity relaxation to a climatological field in the control run. In the runs here we are of course assimilating profile salinity, but the short background correlation length scales mean that salinity increments are only applied locally, and because the data are quite sparse, any strong model drifts are relatively unconstrained by the salinity profile data assimilation. Note also that while NEMOVAR does allow for subsurface cross covariances of temperature and salinity this does not apply at the surface or in the mixed layer; therefore, there is no impact of SST data on salinity increments. The assimilation scheme has recently been developed to include multiple covariance length scales, and by including a longer length scale in future runs, we may be able to constrain the errors more effectively. In the future, also, sea surface salinity data will be available also from SMOS and Aquarius. It is of course still preferable to fix the model bias rather than relying on observations or assimilation improvements to control the problem.

It seems likely that changes in the runoff play a major role in these salinity differences since other aspects of the ocean model, including parameterizations, are the same for the ocean control and the coupled ocean. For the runoff the ocean control uses climatological estimates from Bourdalle-Badie and Treguier (2006) based on Dai and Trenberth (2002), but the coupled model uses the JULES surface and subsurface runoff routed to the ocean using the Total Runoff Integrating Pathways model (TRIP; Oki and Sud 1998). To assess the freshwater input into the ocean, we calculate the evaporation minus precipitation minus runoff values in the region around the river basin (from 38.48°S, 58.5°W to 33.63°S, 53.5°W; see Fig. 6). There is a significant difference between the two runs, as the coupled model inputs significantly more freshwater into the small region of the Rio Plata outflow; additionally, there is a much amplified annual cycle. The evaporation minus precipitation results are relatively similar between the two systems in this region of ocean at, for example, 0.2 × 10^8 kg s^{-1} in December 2011, so we conclude that the coupled model runoff is very different to the climatological runoff in the ocean control. Direct validation of the runoff is not available, but given the drift in salinity, it seems most likely that the coupled model has the more unrealistic value. Why this might be is not clear; precipitation model errors over the river basin are insufficient to explain the difference (not shown) directly. In future coupled DA runs we will consider using climatological...
runoff into the ocean rather than the coupled model value until a better solution is found.

c. Atmosphere and ocean increments

Examining the assimilation increments provides another method for assessing model errors and biases and whether these are improved by coupled data assimilation. Coupled DA increments give direct information about the spatial patterns of short-range forecast errors in the coupled model. The atmosphere surface air temperature mean and standard deviation of the increments for December 2011 are shown in Fig. 7. Over land the coupled and control DA results indicate some significant biases of ~1°C every 6 h. These findings are generally similar for the coupled and control DA cases. Major differences are seen over the Arctic region and subpolar North Atlantic. However, neither the coupled nor the control run is consistently improved. There is an intriguing small reduction in the mean increments in some parts of the equatorial and South Pacific regions. There is a potential issue with the large lakes (e.g., Lake Victoria) and inland seas (e.g., Caspian Sea) in the coupled DA cases where the absolute mean increments are seen to be bigger than in the control. In the ocean control run, a carefully determined lake surface temperature from OSTIA is used (Fiedler et al. 2014). The coupled

FIG. 4. The 0–10-m depth salinity RMS binned innovations (in psu) in $10^8 \times 10^8$ bins over all 13 months for the (a) coupled and (b) ocean control analyses. The triangular marker indicates the location of the Rio de la Plata.
(and ocean control) simulation uses NEMO to model the large lakes and inland seas, but there is currently no ocean data assimilation in the lakes. The modeling and assimilation of data over the lakes clearly needs closer attention in future runs.

In Fig. 8 we show the mean and standard deviation of the SST increments applied to the ocean model. These are very similar for the coupled and control cases. The main features are negative mean increments in the tropical band, suggesting a positive model SST bias. In some parts of the South Pacific there are also strong negative mean increments. If we compare the coupled and control results, we see that the coupled absolute mean increments are smaller in many of these biased regions, an encouraging finding. As mentioned above, there may be reductions in the mean surface air temperature increments in the same regions in the atmosphere (Fig. 7). However, in the South Pacific and other Southern Hemisphere oceans the standard deviation of the increments is increased in the coupled case. This is strongly linked to the increased SST innovations in the coupled case, discussed in section 3b, as a result of the diurnal cycle representation. When a new diurnal model is implemented in the coupled model, the mean and
standard deviation of the coupled DA increments will be a good way to assess the impact of this change.

4. Coupled forecast results

To fully investigate the impact of coupled DA, we now consider the performance for selected forecast periods. Two sets of coupled forecasts were run for selected periods (1–31 December 2011, 26 August–15 September 2012, and 20–31 October 2012); the first starting with separate control ocean and atmosphere analyses (fc_ctl) and a second set using the coupled DA analysis (fc_cpld). The idea is to see if coupled DA offers a more balanced initialization, which may lead to improved coupled model forecast performance.

First, we consider the atmosphere results. Examples for the period 26 August–15 September 2012 are shown in Fig. 9 (left), where model surface air temperature is compared to surface in situ observations (the same as those that are later assimilated) and the results are plotted as a function of forecast lead time. The general pattern is that as the forecast lead time increases, the fit to the observations is degraded. There is however very little difference between the performance of the coupled DA and control DA forecasts. The Northern Hemisphere 1.5-m temperature RMS error in the coupled DA forecasts appears slightly smaller from day 8 onward than in the control DA forecasts, but even in that case the difference does not appear to be significant. For the other forecast periods (December 2011 and October 2012) there are no significant differences in atmosphere performance between the coupled and control initializations.

The ocean forecast observation comparison, performed against in situ SST observations [see Fig. 9 (right) column], again shows little notable difference between the coupled and control DA. The only significant difference between the sets of forecasts is in the Northern Hemisphere temperature bias in August–September 2012, with the coupled forecasts showing a bias that is smaller and of the opposite sign to the control forecasts. Additional ocean forecast results for SLA and profile temperature and salinity are shown in Fig. 10. For December 2011 the coupled and control initialized forecast results are almost identical for these variables and also for SST (not shown for December 2011). In contrast, in August–September 2012 the coupled DA initialized forecasts are degraded slightly (and barely significantly) in SLA compared to the control, but more significantly in the profile salinity. Temperature and SST forecasts still perform equally well for both sets of forecasts, however. In section 3b we learned that the salinity was degrading with time in the coupled analysis because of problems with the runoff. It seems likely that fixing the runoff problem in the coupled DA as described in this section will resolve the issue in the forecasts with salinity.
5. Discussion and conclusions

The work presented here demonstrates the first Met Office implementation of a coupled DA system. To test the performance of coupled DA, we performed separate (“control”) ocean and atmosphere DA analyses over a 13-month period, and compared the results to our coupled DA analysis. The model parameters and resolution were as far as possible identical for the coupled and control systems. A number of coupled model forecasts were initialized from the coupled DA analysis and from the control analyses, the latter forecasts being representative of the way that the current GloSea 5 seasonal forecast system (MacLachlan et al. 2015) is initialized.

The coupled DA generally performed reasonably well. The results for coupled DA were mostly very similar to the uncoupled DA. This is encouraging considering that the coupled DA system is new and neither

FIG. 7. (a),(c),(e) Mean and (b),(d),(f) standard deviation of surface air temperature increments [°C (6 h)⁻¹] for December 2011. Shown are the results for the (a),(b) coupled and (c),(d) control simulations. (e) The coupled and control difference in the absolute mean and (f) the coupled and control difference in the standard deviation of the increments. In the difference plots blue implies that the coupled DA has smaller increments.
the ocean DA nor atmosphere DA systems have been tuned or changed in any way for coupled DA. In particular, no specific coupled DA tuning of the error covariances has yet been done.

Although the overall performance of the coupled DA was reasonable, there were some problems identified. One was an issue with the diurnal cycle in the ocean compared to in situ SST observations. This is a general difficulty with using an ocean model with a 1-m top level without an explicit diurnal skin model. The coupled ocean model has an amplified diurnal cycle compared to the uncoupled ocean, so the problem is more evident in that case. Work is on going to include an explicit diurnal model in the ocean, which should reduce this error. Coupled DA will provide a good way to assess the performance of this development. The other main issue was with the coupled river runoff, which appeared to be in error; it is unclear at this stage whether this is due to errors in the model or the DA system. Over time this led to a degradation of the salinity around some river basins,
which also degraded the SLA in those regions. As an interim measure, for future coupled DA runs we may consider using climatology to give the ocean river inflow while improvements to the system are developed. Another aspect that may be improved is the assimilation and modeling of large lakes and inland seas. In the coupled DA case this did seem to cause nearby local degradations in the atmosphere results.

A particular benefit of coupled DA is that it directly challenges the coupled model with observations highlighting short-term model biases. Using comparisons of observations and the model background and analyses, as well as the mean assimilation increments, starts the work of diagnosing sources of error close to the time at which these originally develop. Reducing these errors will not only improve the quality of short-range coupled forecasts, but can feed back into improved coupled model performance on longer time scales, such as those for seasonal forecasting and climate prediction. This extends the Met Office’s seamless approach to weather and climate prediction and model development (Brown et al. 2012) from atmospheric to fully coupled modeling.

Using what we have learned from these experiments, the next step is to develop an operational coupled short-range forecasting system, which can be compared in near–real time to the existing separate NWP and ocean forecasting systems and ultimately, if the comparison is favorable, replace them. Meanwhile, other work will be

![Fig. 9. Surface air temperature error against forecast lead time calculated from model–observation differences (RMS differences are shown as solid lines and mean differences are shown as dotted lines) for the (a) Northern Hemisphere, (c) tropics, and (e) Southern Hemisphere 1.5-m surface temperature error (°C; for the August–September 2012 period). SST forecast error (°C) for the same period in the (b) Northern Hemisphere, (d) tropics, and (f) Southern Hemisphere. The bars indicate the 68% confidence limits.](image-url)
to develop more strongly coupled DA (i.e., more coupling between the data assimilation systems). The benefits of this are to extract the maximum information from the ocean and atmosphere for observations near their interface. As with the existing separate assimilation systems, it is still likely that the increments are not consistent between the two systems in the weakly coupled DA and therefore not well balanced. There is some evidence in our system, contrary to expectations and also contrary to a study of the ECMWF coupled DA system by Mulholland et al. (2015), that both the uncoupled and weakly coupled DA initializations generate similar initialization shocks (looking at, for example, gravity waves in 20°C isotherm depth) (results not shown). In experiments with simplified coupled systems, weakly coupled DA systems often still show significant initialization shocks (Smith et al. 2015). Strongly coupled DA, by combining the ocean and atmosphere DA systems and ensuring that the increments are balanced, improves this situation. One step toward strongly coupled DA is to learn more about the covariances between the ocean and the atmosphere. This will help to quantify the potential benefits of moving to strongly coupled DA and will give the coupled covariances needed for any prototype strongly coupled DA system. In the shorter term, work of improving the error covariances of the ocean and atmosphere separately, but within a coupled context, should also help to improve the performance of the weakly coupled DA.
Acknowledgments. This work was supported by the European Commission MyOcean2 project and the U.K. Royal Navy.

REFERENCES


Laloyaux, P., M. Balmaseda, K. Mogensen, P. Janssen, and D. Dee, 2014: The ECMWF coupled assimilation system. Proc. 2014 Conf. on Big Data from Space (BiDS’14), Frascati, Italy, European Space Agency-European Space Research Institute, 16–19. [Available online at http://dx.doi.org/10.2788/1823.]


Mogensen, K. S., M. A. Balmaseda, A. Weaver, M. Martin, and A. Vidard, 2009: NEMOVAR: A variational data assimilation system for the NEMO ocean model. ECMWF Newsletter, No. 120, ECMWF, Reading, United Kingdom, 17–21.


